



AHRQ QUALITY INDICATORS

Patient Safety Indicators

SOFTWARE DOCUMENTATION



AHRQ Quality Indicators

**Patient Safety Indicators:
SPSS Software Documentation, Version 2.1**

Department of Health and Human Services
Agency for Healthcare Research and Quality
www.ahrq.gov

May 28, 2003
AHRQ Pub. No. 03-R205

Citation

AHRQ Quality Indicators -- Patient Safety Indicators: Software Documentation, Version 2.1 - SPSS. Rockville, MD: Agency for Healthcare Research and Quality, 2003. AHRQ Pub.03-R205

Abstract

The value of information on health care quality has never been so widely recognized; yet many organizations lack the resources and/or expertise to build a quality information program from the ground up. Recognizing this, the AHRQ Quality Indicators were developed specifically to meet the short-term needs for information on health care quality using standardized, user-friendly methods and existing sources of data.

This module of the AHRQ Quality Indicators was designed to capitalize on the availability of administrative data on inpatient stays to produce information for Patient Safety Indicators (PSIs). These indicators capture potentially preventable complications and iatrogenic events for patients treated in hospitals. They can be used as a screening tool for problems that patients experience as a result of exposure to the healthcare system and that are likely amenable to prevention by changes at the system or hospital level.

This document is the software documentation for the Patient Safety Indicators software Version 2.1, which is provided on the AHRQ Web site. The software was developed in SPSS, for use on a personal computer. By making this tool available, we hope to assist others in producing information on health care quality more cost effectively.

Details on the development of the Inpatient Quality Indicators can be found in "Guide to Patient Safety Indicators."

Acknowledgments

This product is based on the work of many individuals who contributed to its development and testing.

The following staff from the **Evidence-based Practice Center (EPC) at UCSF-Stanford** performed the evidence review, facilitated clinician reviews, completed the empirical evaluation, and created the programming code and technical documentation for the new Patient Safety Indicators:

Core Project Team

Kathryn M. McDonald, M.M. (Stanford)
Jeffrey Geppert, J.D. (Stanford)
Sheryl M. Davies, M.A. (Stanford)
Patrick Romano, M.D., M.P.H. (UC-Davis)

Other Contributors

Kaveh G. Shojania, M.D. (UCSF)
Bradford Duncan M.D. (Stanford)
Angela Hansen, B.A. (Stanford)

The following staff from **Social & Scientific Systems, Inc.**, developed this software product, documentation, and guide:

Programmers

Leif Karell
Kathy McMillan
Fred Rohde

Technical Writer

Patricia Burgess

Graphics Designer

Laura Spofford

Contributors from the **Agency for Healthcare Research and Quality**:

Anne Elixhauser, Ph.D.
Marlene Miller, M.D., M.Sc.
Denise Remus, Ph.D., R.N.

H. Joanna Jiang, Ph.D.
Margaret Coopey, R.N., M.G.A, M.P.S.

We also wish to acknowledge the invaluable contribution of the clinical panelists and the peer reviewers of the evidence report; and the beta-testers of the software products.

Table of Contents

Abstract	i
Acknowledgments	ii
Introduction	1
Components of the Patient Safety Indicators module	1
Quick Reference	2
Figure 1. Processing Steps for the Patient Safety <u>Hospital</u> Indicators	3
Figure 2. Processing Steps for the Patient Safety <u>Area</u> Indicators	4
Patient Safety Indicator Module Details	5
Background	7
Data elements and coding conventions	8
Input for the AHRQ Quality Indicators	8
Coding of diagnoses, procedures, and DRGs	9
Use of External Cause of Injury Codes	10
Missing values	10
User options	13
Processing steps	14
1. Identify outcomes in inpatient records	14
2. Identify populations at risk	14
3. Calculate observed (raw) Patient Safety Indicator rates.....	14
4. Risk adjust the Patient Safety Indicator rates. Hospital PSIs only.....	15
5. Create MSX smoothed rates. Hospital PSIs only.	15
Program descriptions	16
Program 1: run PSSPS1	16
Program 2: run PSSPSP2.....	17
Program 3: run PSSPSP3.....	19
Program 6: run PSSPSA2.....	21
Reviewing the printed output.	23
PSSPS1.....	23
PSSPSP2	24
PSSPSP3	24
PSSPSA2	26
Interpreting the results	27
Observed means.....	28
Risk-adjusted rates. Hospital PSIs only.	28
MSX smoothed rates. Hospital PSIs only.	29
Comparing observed, risk-adjusted and smoothed rates. Hospital PSIs only.....	30
Confidence Intervals.....	31
Benchmark timings	31
User Support and Information	32

List of Tables

Table 1. Patient Safety Indicator (PSI) Variables	5
Table 2. Prefixes for the Patient Safety Indicator (PSI) Variables	6
Table 3. Patient Safety Module Contents.....	6
Table 4. Data Elements and Coding Conventions	11
Table 5. PSSPSP2 aggregation possibilities.....	18
Table 6. PSSPSA2 aggregation possibilities.....	22

AHRQ Quality Indicators, SPSS Software Documentation, Version 2.1: Patient Safety Indicators

Introduction

This documentation describes the software for implementing the Patient Safety Indicator (PSI) module of the AHRQ Quality Indicators Version 2.1 and highlights features of the analytic approach of particular interest to new users. Detailed information about the Patient Safety Indicators (PSI) is contained in the companion document “*Guide to Patient Safety Indicators*”, which includes information about the development of the indicators and their definitions.

The software consists of a series of computer programs that:

- ▶ Assign and calculate PSIs from hospital discharge abstracts or UB (uniform billing) claims data.
- ▶ Print PSI results in SPSS format which can be converted to a word processing, spreadsheet, or graphics format, at the user's option.
- ▶ Create risk-adjusted rates that adjust for case mix differences.
- ▶ Create smoothed (i.e., reliability-adjusted) rates that reduce fluctuations over time due to random variation.

This documentation includes:

- ▶ A discussion of the data elements necessary to produce the AHRQ PSIs and the uniform coding conventions recommended for those elements.
- ▶ Descriptions of the PSI processing steps in non-technical language.
- ▶ Detailed descriptions of the functions of each PSI SPSS program.

Critical user information is highlighted with this symbol: 

Components of the Patient Safety Indicators module

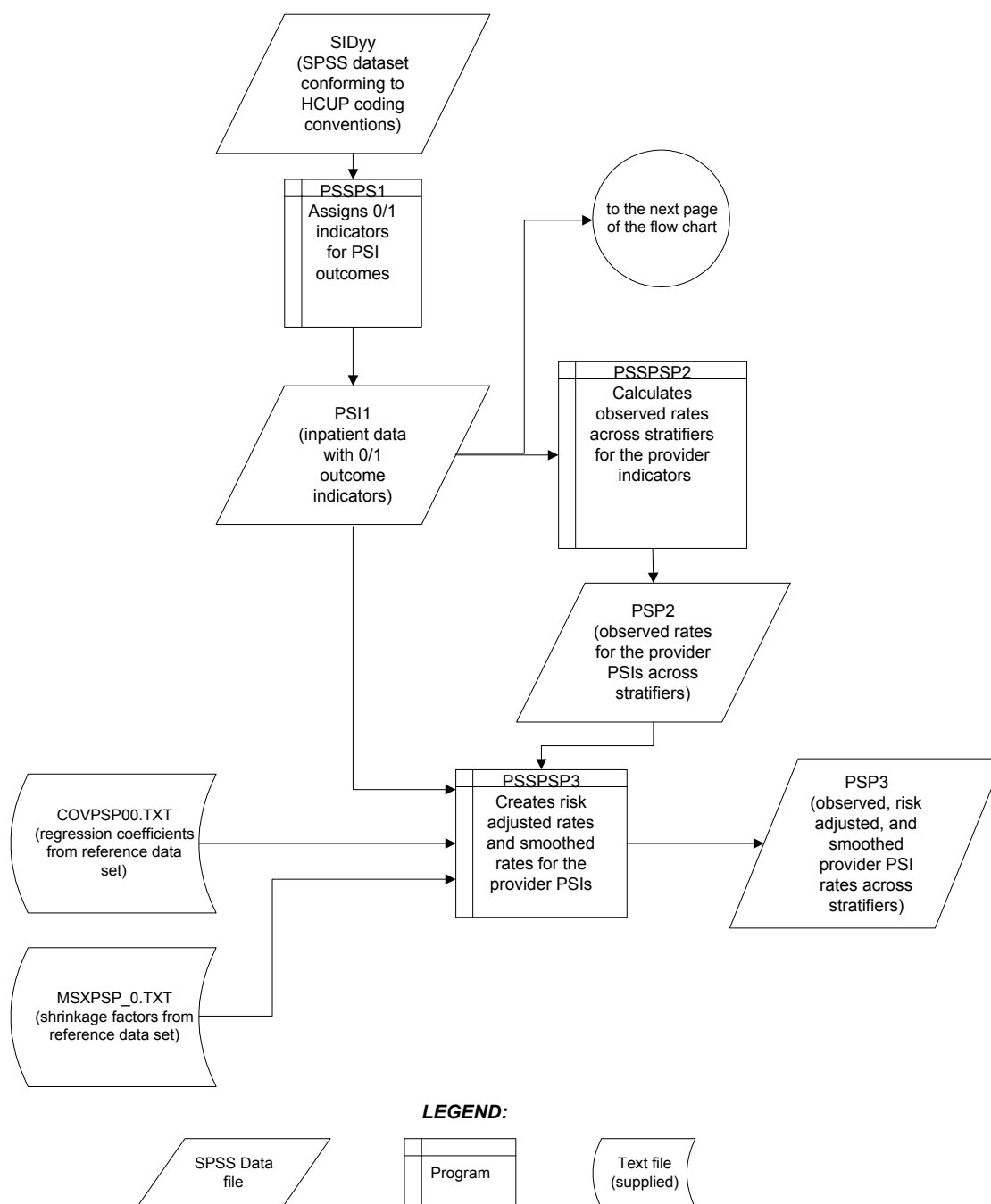
As shown in Table 3 on page 6, the Patient Safety Indicators module consists of four SPSS programs and six ASCII text files that contain auxiliary data. These programs and text files are described in the subsequent sections of this document. The programs were developed for use in Statistical Package for the Social Sciences (SPSS) version 10.0 or higher, on an IBM compatible computer.

A note on the naming conventions for the Patient Safety Indicators (PSI) module programs. Most programs have names of the form PSSPSi, PSSPSAi, or PSSPSPi. The first two characters “PS” of the program name indicate a Patient Safety Indicators program, to distinguish it from other modules that have been released from AHRQ. The next three characters of the program name are “SPS” and are present to distinguish the program from the SAS versions of the software. In the programs that are specific to either Area or Hospital runs, the sixth character of the program name is either “A” or “P”. The sixth character is an “A” if it is for the production of Area indicator rates that use county and Metropolitan Statistical Area (MSA) area populations as denominators. The sixth character is a “P” if it is a program that is for the production of Hospital indicator rates that use subsets of the hospitalized population. The last character (i) of the variable name designates the number of the specific program.

Quick Reference

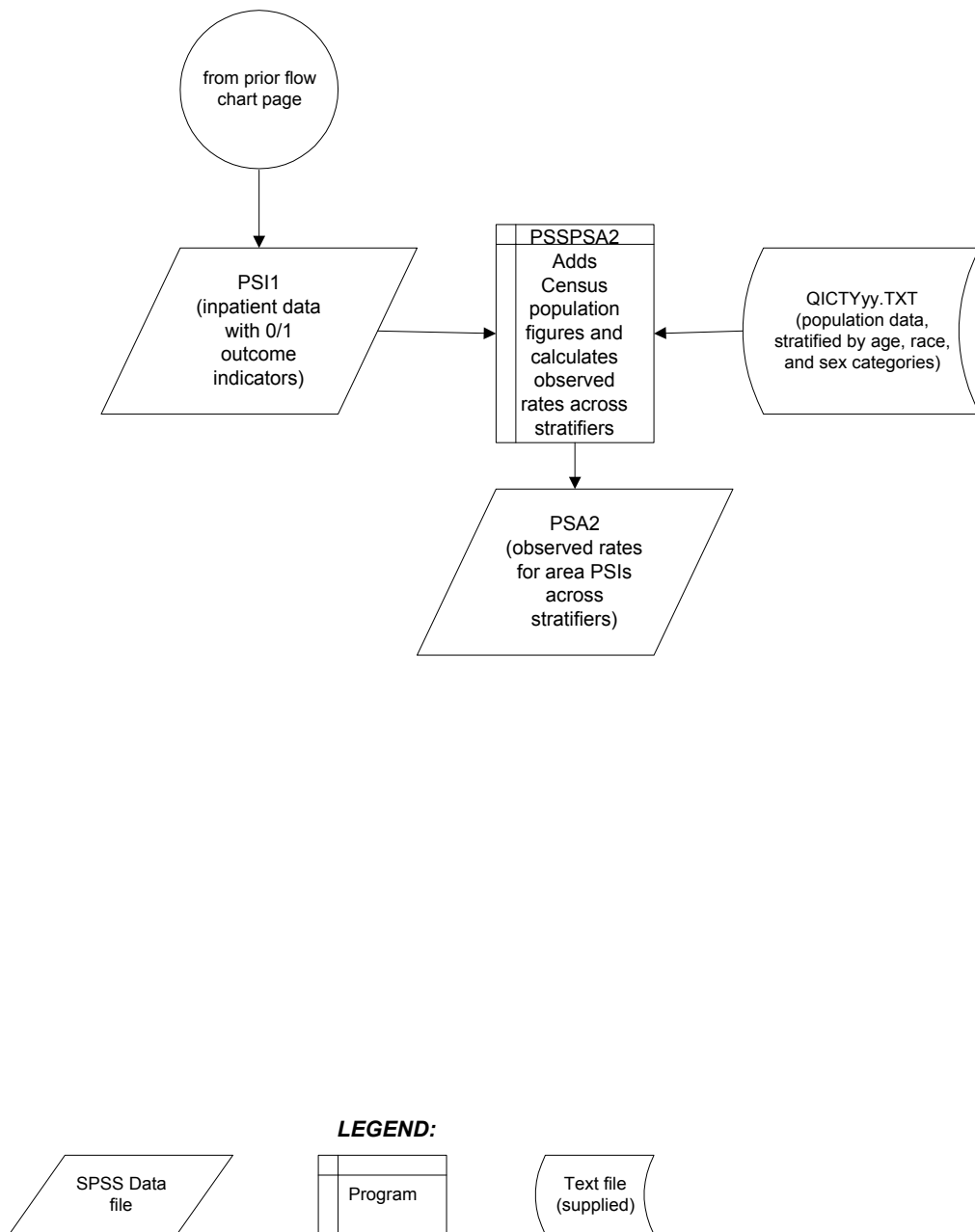
The subsequent four pages are intended to serve as a quick reference to assist in reading this documentation and in reviewing the Patient Safety Indicator (PSI) module outputs. Processing steps (Figures 1 and 2) are shown first followed by a listing of the module indicator variables (Table 1), variable prefixes (Table 2), and module contents (Table 3). One suggestion is to print these pages in duplex mode on two sheets of paper so that they can then be easily referred to as the need arises.

Figure 1. Processing Steps for the Patient Safety Hospital Indicators



Note: There is another version of PSSPS3 that is available for calculating confidence intervals around each Patient Safety Indicator. This syntax, PSSPS3CI is available upon request. Please send an email to QI support at support@qualityindicators.ahrq.gov with "PSI Confidence Interval Syntax Request" as the subject line and indicate in the body of the e-mail you would like the SPSS syntax.

Figure 2. Processing Steps for the Patient Safety Area Indicators



Patient Safety Indicator Module Details

Table 1. Patient Safety Indicator (PSI) Variables		
Type	PSI number	Indicator
Hospital	1	Complications of anesthesia
	2	Death in low mortality DRGs ^a
	3	Decubitus ulcer
	4	Failure to rescue
	5	Foreign body left during procedure
	6	Iatrogenic pneumothorax
	7	Selected infections due to medical care
	8	Post-operative hip fracture
	9	Post-operative hemorrhage or hematoma
	10	Post-operative physiologic and metabolic derangements
	11	Post-operative respiratory failure
	12	Post-operative pulmonary embolism or deep vein thrombosis
	13	Post-operative sepsis
	14	Post-operative wound dehiscence
	15	Accidental puncture or laceration
	16	Transfusion reaction
	17	Birth trauma
	18	Obstetric trauma – vaginal with instrument
	19	Obstetric trauma – vaginal without instrument
	20	Obstetric trauma – cesarean section
Area	21	Foreign body left during procedure
	22	Iatrogenic pneumothorax
	23	Selected infections due to medical care
	24	Post-operative wound dehiscence
	25	Accidental puncture or laceration
	26	Transfusion reaction

^a The second indicator is reported as a single measure, but is also stratified into medical (adult and pediatric), surgical (adult and pediatric), obstetric and psychiatric DRGs.

Table 2. Prefixes for the Patient Safety Indicator (PSI) Variables			
Prefix	Contents	Examples	
		Hospital indicator # 8, Post-operative Hip Fracture	Area indicator # 23, Infection Due to Medical Care
T	Inpatient numerator (top)	TPPS08	TAPS23
P	Population denominator (pop)	PPPS08	PAPS23
O	Observed Rate	OPPS08	OAPS23
R	Risk-adjusted Rate	RPPS08	RAPS23
S	Smoothed Rate	SPPS08	SAPS23

Table 3. Patient Safety Module Contents	
SPSS programs	ASCII text files †
PSSPS1.SPS	QICTY97.TXT
PSSPSP2.SPS	QICTY98.TXT
PSSPSP3.SPS	QICTY99.TXT
PSSPSA2.SPS	QICTY00.TXT
	QICTY01.TXT
	COVPSP00.TXT
	MSXPSP_0.TXT

† The ASCII text files are data files provided with the Patient Safety module and are necessary for the programs to run. There is no need to convert these files to SPSS format for use with the software.

Background

The Patient Safety Indicator module contains 26 indicators that reflect the quality of care inside hospitals. These indicators serve as flags for potential quality problems rather than definitive measures of quality of care. The indicators include two groups of measures based on hospital stays.

1. There are 20 hospital-based patient safety indicators for medical conditions and surgical procedures that have been shown to have complication/adverse event rates that vary substantially across institutions and for which evidence suggests that high complication/adverse event rates may be associated with deficiencies in the quality of care. These indicators are measured as rates -- number of complications / adverse events divided by the number of admissions for the procedure or condition. The hospital-based indicators include only those cases where a secondary diagnosis code flags a potentially preventable complication.

Eight of these indicators are for surgical discharges, eight are for either medical or surgical discharges, and four are for obstetric discharges. Two of the hospital-based patient safety indicators, Post-operative hemorrhage or hematoma (#9) and Post-operative pulmonary embolism or deep vein thrombosis (#12), will incorporate information about when procedures were performed (relative to the admission date) if this information is available. Note, however, that if the day of procedure information is missing, the computed rates for post-operative hemorrhage or hematoma will be slightly higher and the computed rates for post-operative pulmonary embolism or deep vein thrombosis will be slight lower than if the information was available. Admission type is used by four of the PSIs: Post-operative physiologic and metabolic derangements, Post-operative respiratory failure, Post-operative sepsis, and Birth trauma (#s 10, 11, 13, and 17), to identify elective surgeries and newborn admissions.

2. There are six area-based patient safety indicators, where geographic areas can be defined at the MSA (Metropolitan Statistical Area) level for urban areas and at the county level for rural areas or by county for all areas. MSAs are defined in terms of entire counties, which include the county (or counties) that contain the largest city and any adjacent counties that have at least 50 percent of their population in the urbanized area surrounding the largest city. In New England, the MSAs are defined in terms of cities and towns rather than counties. The MSA definitions in this software are based on the standards effective June 30, 1999.

These indicators are identical to hospital-based indicators, except that the numerator uses principal diagnosis in addition to secondary diagnoses in order to capture all cases of the complication, and not only those that develop during a given hospital admission.

The 26 indicators in the PSI module are listed in Table 1 on page 5. See also Appendix A of the "*Guide to Patient Safety Indicators*" for the definition of each indicator. Note that the second indicator (for death in low mortality DRGs) is reported as a single measure, but also has seven sub-indicators for medical (adult and pediatric), surgical (adult and pediatric), obstetric and psychiatric DRGs.

The data required for measuring these indicators come from hospital discharge abstracts or billing claims (administrative data) which are readily available in many states. The residential population data for the six area-based utilization indicators are from the Bureau of Census.






The software generates observed rates for the area-based patient safety indicators and observed, risk-adjusted, and smoothed (i.e., reliability-adjusted) rates for all hospital based indicators. Observed rates are the raw rates. Risk-adjusted rates are derived from applying the average case mix of a baseline file that reflects a large proportion of the U.S. hospitalized population. Smoothed rates are estimates with removal of fluctuations over time due to random variation. The observed rates for the area PSI are extremely low, and as a result age-sex adjustments are not necessary.

For information about how these indicators were identified, see the "*Guide to the Patient Safety Indicators*."

Data elements and coding conventions



Input for the AHRQ Quality Indicators

The PSI software was written to process data from discharge data abstracts that contain information about hospital stays. The specific data elements that are needed for the PSI software are listed in Table 4. The PSI module was tested on data from the Healthcare Cost and Utilization Project (HCUP). HCUP is an ongoing Federal-State-private collaboration to build uniform databases from administrative hospital-based data collected by State data organizations and hospital associations.


-  *The input data files for the Patient Safety Indicators software must be in SPSS format.*
-  *We recommend that users recode data elements in their input files to be consistent with the coding expected by the software. This will minimize the number of internal changes that will be necessary in the Patient Safety Indicators software. For example, recoding the SEX data element in the input file to be consistent with the coding described in Table 4 (i.e., 1 for male and 2 for female) is easier than modifying all uses of the Sex data element in the PSI programs.*
-  *Not every value for every data element is referenced by the PSI software. For example, admission source (ASOURCE) values are used to identify transfers. However, we include the complete set of HCUP values to assist users of the uniform HCUP data files.*
-  *Four of the PSIs (#s 10, 11, 13, and 17) use admission type (ATYPE) to identify elective surgeries and newborn admissions. If admission type (ATYPE) is not available in the user's data, the user may want to evaluate other data in the file to create a proxy to identify emergency, urgent and newborn admission types.*
-  *To minimize internal changes to the software, all required elements should be present in the input data file. If a required element is not available, a dummy element should be provided. Failure to provide a dummy element will result in errors during the execution of the PSI programs.*


The data elements listed in Table 4 are those required for the input files of both SAS and SPSS. When a variable is not required for the program the notation "Not used by the QI program" will be present in the Comments column. The SAS program will still run if these variables are not present, however, the SPSS program will not run correctly. Standardizing the variables and data values in the input data file will be helpful for users who have access to both statistical software packages.

The data element HOSPSTCO provides flexibility for users to calculate the area PSIs by hospital location or by patient residence. If the user wants to calculate the PSIs using the population associated with the hospital location as the denominator, the values for this variable should be the individual hospital FIPS¹ state/county codes. If the user wants to calculate the area PSIs based on the population of the MSA region or county associated with the patient residence, the values for this variable should be the FIPS state/county code associated with each individual patient's place of residence.


-  *If the hospital FIPS code is used in HOSPSTCO, users should be aware that rates may be biased for hospitals which serve as regional referral centers. These hospitals are likely to treat patients from outside the MSA, county or even the state in which the facility is located.*
-  *Using the patient FIPS state/county code for analysis may more accurately reflect the true population at risk.*

¹ Federal Information Processing Standard, as defined by the U.S. Department of Commerce, National Institute of Standards and Technology (formerly National Bureau of Standards).

 It is possible that some records in the input data file may be missing the patient FIPS code. Users should be aware that any records with missing values (in the HOSPSTCO data field) are excluded from the calculations of observed, risk-adjusted and smoothed PSI area rates. They will be included in the output from the first program (PSSPS1).


 The software will generate area PSI rates for each county included in the HOSPSTCO data field. Users are encouraged to limit cases in their input file to those patients discharged from the geographic area of interest. For example, if you are using data from the state of Wisconsin and are interested in generating a rate for Wisconsin, you should remove the cases where the patient's county of residence (FIPS code) is from another state. Otherwise the total or overall rate will include the population of the counties outside Wisconsin.

The software provides the user with the option of producing output by MSA or by county. When MSA is selected, urban areas are always defined by MSAs. When county is selected urban areas will be defined by county. Rural areas are always defined by county.


 In the software programs, the MSALEVL parameter indicates whether MSA codes should be created from the FIPS state/county codes in the input dataset. If the user wishes to analyze data at the MSA level, this parameter should be set to "Yes". If the user wishes to analyze data at the county level, this parameter should be set to "No".

Coding of diagnoses, procedures, and DRGs

Diagnoses and procedures must be coded using the International Classification of Diseases, 9th Revision, Clinical Modification (ICD-9-CM)

 Significant modifications to ICD-9-CM codes occurred in the early 1990s. PSI definitions only reflect codes valid after October 1, 1997, and therefore may not accurately analyze data collected before 1997.

- ◆ Diagnoses and procedure codes should be in character strings.
- ◆ Diagnosis and procedure codes should not contain any decimal points.
- ◆ Diagnosis and procedure codes should be left justified. Diagnosis codes with fewer than five digits should be padded with spaces (not zeros) on the right. Procedure codes with fewer than four digits should be padded with spaces (not zeros) on the right.


 The PSI programs assume accurate and specific coding of ICD codes. If an ICD code is specified using 5 digits, the software recognizes only 5-digit versions of that code and not truncated versions at 3 or 4 digits. Similarly, if an ICD code is specified using 4 digits, the software recognizes only 4-digit versions of that code and not truncated versions at 3 digits. For example, 4281 is the ICD-9-CM diagnosis code for left heart failure. If your data are coded less specifically (i.e., 428 is used to designate "heart failure"), these less specific codes are not recognized by the software and will result in undercounting cases.

- ◆ This software has ICD-9-CM codes updated through FY2003 changes.
- ◆ Diagnosis-related groups (DRGs) are those derived from the Centers for Medicare & Medicaid Services (previously Health Care Financing Administration) Medicare grouper. The software expects that you will be using the DRG effective on the discharge date. The software now refers to data elements DRG and MDC. Your data should be coded accordingly. The software may be modified at the user's option to use other types of DRGs. However, the impact of using other types of DRGs should be evaluated carefully before making such a change.

Use of External Cause of Injury Codes

External cause of injury codes - E-codes – are used to classify environmental events, circumstances, and conditions as the cause of injury, poisoning, or other adverse events. The use of E-codes is not always required by a state uniform billing committee or a state data organization. Users should be knowledgeable of the E-code requirements and practices of hospitals represented in the input data file. Several of the indicators use E-codes in their numerator or denominator definitions. See the “*Guide to the Patient Safety Indicators*” for additional information.

Missing values

 The PSI programs do not distinguish among different types of missing values.

Data files of hospital discharge abstract data may have numeric data elements coded using SPSS system missing "dot" values. For example, besides the standard SPSS value of "." for missing data, there might also be present values of "-1" for invalid data, "-2" for data unavailable from a particular source, "-3" for inconsistent data. However, the PSI programs do **not** distinguish among the different types of missing codes. Therefore, all types of missing values in the input data to the PSI module can be represented by a single value - missing data (.) for numeric variables and blank (' ') for alphanumeric (or character) variables.


Table 4. Data Elements and Coding Conventions

Variable name	Description	Format	Value description	Comments
KEY	Sequence number. Unique case identifier.	Numeric	User defined unique numeric identifier for each discharge record	Not used by the PSI programs, but should be present to facilitate data exploration and possible trouble shooting.
AGE	Age in years at admission	Numeric	Age in years	If missing, excluded from the analysis.
RACE	Race of patient.	Numeric 1 2 3 4 5 6	White Black Hispanic Asian or Pacific Island Native American Other	The values 1,2, and 3 are used in the QI software for stratification (if selected by user). All other ethnicity codes are mapped to an 'other' category.
SEX	Sex of patient	Numeric 1 2	Male Female	If missing, excluded from the analysis.
PAY1	Expected primary payer	Numeric 1 2 3 4 5 6	Medicare Medicaid Private, incl. HMO Self-pay No charge Other	The values 1,2,3, and 4 are used in the QI software for stratification of the hospital-level PSIs (if selected by user). All other payer codes are mapped to an 'other' category.
HOSPSTCO	Hospital location or location of patient residence (FIPS State/county code)	Numeric ssccc	Modified Federal Information Processing Standards State/county code	Available at: http://eire.census.gov/popest/archives/files/fips.txt
HOSPID	Data Source hospital number	Numeric	Hospital identification number	Used to facilitate data exploration and possible trouble shooting.
DISP	Disposition of patient	Numeric 1 2 3 4 5 6 7 20	Routine Short-term hospital Skilled nursing facility Intermediate care Another type of facility Home health care Against medical advice Died in the hospital	The values 2 and 20 are referenced by the PSI code (to identify transfers to another short-term hospital and patients who died in the hospital).
ATYPE	Admission Type	Numeric 1 2 3 4 5 6	Emergency Urgent Elective Newborn Delivery Other	The values 3 and 4 are referenced by the PSI code (to identify elective surgeries and newborn admissions). PSIs 10, 11, 13, and 17 will be affected if ATYPE values are missing.
ASOURCE	Admission Source	Numeric 1 2 3 4 5	ER Another hospital Another facility incl LTC Court/law enforcement Routine/birth/other	The values 2 and 3 are referenced by the PSI code (to look for transfers from another hospital or facility)
LOS	Length of Stay	Numeric	Number of days from admission to discharge	
DRG	Diagnosis Related Group	Numeric	DRG from CMS DRG Grouper	

Variable name	Description	Format	Value description	Comments
MDC	Major Diagnostic Category	Numeric	MDC from CMS DRG grouper	
DX1 – DX30	Diagnoses ICD-9-CM codes. DX1 is the principal diagnosis, DX2-DX30 are secondary diagnoses. Note: If e-codes are separated from secondary diagnoses in the input data file, the variable should be renamed and included as a secondary diagnosis variable (e.g., e-codes would be labeled as DX10 in a data file where the last secondary dx field is DX9).	String, 5 character	Diagnosis codes	Users must specify the number of diagnosis variables as the parameter in the PSSPS1.SPS program.
NDX	Count of the number of diagnoses (Dx) on the discharge record.	Numeric	Count of Diagnoses	If the 1 st discharge record has 5 Dx and the 2 nd has 10, NDX would be 5 on the 1 st and 10 on the 2 nd . Not used by the SPSS program.
PR1 – PR30	Procedure ICD-9-CM codes. PR1 is the principal procedure, PR2-PR30 are secondary procedures.	String, 4 character	Procedure codes	Users must specify the number of procedure variables as a parameter in the PSSPS1.SPS program.
NPR	Count of the number of procedures (Pr) on the discharge record.	Numeric	Count of Procedures	If the 1 st discharge record has 1 Pr and the 2 nd has 2, NPR would be 1 for the 1 st record and 2 for the 2 nd . Not used by the SPSS program.
PRDAY1 – PRDAY30	Days from admission to procedure. PR1 is the principal procedure, PR2-PR30 are secondary procedures.	Numeric	Days from Admission to Procedure	Necessary variables if the user sets the &PRDAY parameter in PSSPS1.SPS to equal one. In this case, it is expected that the number of PRDAY variables agrees with the number of Procedure codes present. These variables are not needed if the user sets the &PRDAY parameter in PSSPS1.SPS to equal zero, indicating that there is no procedure day information available.

User options

The PSI software reflects the development and implementation of the software using data available to AHRQ. Our goal was to develop the tools, illustrate their use, and then encourage others to adopt and use the tools for their own applications. As a result, we expect and encourage users of the software to consider whether and how to modify the PSI software to better serve their local conditions or interests.

 *Modifications to the definitions of outcomes of interest (numerators) or populations at risk (denominators) are possible but not desirable. Maintaining consistent definitions is important. Once definitions are altered, the ability to make comparisons of PSIs based on original definitions is lost. We encourage users to identify ways to improve the PSI methodology and to share their suggestions with us for future updates.*

AHRQ is planning to provide national PSI estimates through its web-based query system, HCUPnet (<http://hcup.ahrq.gov/HCUPnet.asp>). If users change definitions of the PSIs, it will not be possible to compare users' results to the national estimates in HCUPnet.

Processing steps

Each Patient Safety Indicator (PSI) expressed as a rate, is simply defined as:

Outcome of interest / Population at risk

(or)

Numerator / Denominator

Conceptually, five steps are necessary to produce the PSI rates. The following describes the steps and how the software performs them.

1. **Identify outcomes in inpatient records**

Inpatient records are marked to indicate whether they contain the outcome of interest (numerator or “top”) for each of the AHRQ PSI measures.

This is done by setting a series of flag variables, each of which corresponds to the numerator for a particular PSI. For example, if the inpatient record meets the conditions for inclusion in the outcome for PSI #1, then the outcome indicator for PSI #1 is set to one.

This step requires one pass through the discharge-level data and outputs a discharge-level data file containing the flag indicator variables for the outcomes for each PSI.


2. **Identify populations at risk**

For the Hospital PSIs, the populations at risk (the denominators for calculating the PSI rates, or “pop”) are derived from the hospital discharge records.

For the Area PSIs, the populations at risk (the denominators for calculating the PSI rates) are derived from Census population figures.

3. **Calculate observed (raw) Patient Safety Indicator rates**

Using the output data from step 1 and either hospital discharge or Census population data from step 2, the PSI rates are calculated for user-specified combinations of stratifiers. These stratifiers include hospitals, age groups, race/ethnicity categories, sex, and payer categories for the Hospital PSIs, and areas (MSAs/counties), age groups, race/ethnicity categories, and sex for the Area PSIs.

 *The programs calculate observed PSI rates regardless of the number of cases available. However, PSI rates based on only a few cases should be interpreted with caution.*


In work at AHRQ, we do not report rates based on fewer than 30 cases in the numerator or the denominator. This exclusion rule serves two purposes:

- It eliminates unstable estimates based on too few cases.
- It helps protect the identities of hospitals and patients.

This step outputs the observed rates for user-specified categories, alone or in combination. For example, Hospital PSI observed rates could be output at the hospital level, at the payer level, or at the hospital level by payer.

4. Risk adjust the Patient Safety Indicator rates. Hospital PSIs only.

Regression coefficients from a baseline database (reflecting a large proportion of the U.S. population) are applied to the observed rates to risk-adjust the rates observed in the user's data. These baseline file regression coefficients are provided as part of the PSI module. The risk-adjusted rates will then reflect the age, sex, DRG and comorbidity distribution of data in the baseline file rather than the distribution for the hospitals present in the users' data. This will allow risk-adjusted rates produced by various users to be compared directly to one another. The regression coefficients were derived from AHRQ's State Inpatient Databases (SID) for 29 states. The code to generate these baseline coefficients are not part of the PSI module.

 *For the hospital indicators, if there are fewer than three discharges for the particular PSI within a stratification cell, the risk-adjusted rate for that combination of stratifiers is set to missing.*

5. Create MSX smoothed rates. Hospital PSIs only.

Shrinkage factors are applied to the risk-adjusted rates for each PSI in a process called multivariate signal extraction (MSX). These shrinkage factors were calculated from the SID database of 29 states. For each PSI, the shrinkage estimate reflects a 'reliability adjustment' unique to each indicator. The less reliable the PSI over time and across hospitals, the more the estimate 'shrinks' the PSI toward the overall mean. The resulting rate will appear "smoother" than the raw rate, meaning the year-to-year fluctuations in performance are likely to be reduced. More information on interpreting smoothed rates is contained in the interpretation section of this document. The shrinkage factors are provided as part of the PSI software and do not need to be calculated by users.

These five steps reflect the PSI module production in a nutshell. The next section of this document describes the specifics of each software component of the PSI module software.

Program descriptions

This section describes the four SPSS programs that assign, calculate, and print the Patient Safety Indicators.

For each program there is a description, a list of input and output files, and an explanation of changes to the program that may be required. The flow of data through the PSI module programs is shown in the flowcharts in Figures 1 and 2 that begin on page 3.

If you desire to create and examine observed PSI rates, then the PSSPS1 programs will need to be used, followed by PSSPSP2 (for Hospital observed rates) and/or PSSPSA2 (for Area observed rates). If you also wish to create Hospital risk-adjusted and smoothed PSI rates, then you will also need to run the PSSPSP3 program. Risk-adjusted and smoothed rates are not calculated for the Area PSIs.

Program 1: run PSSPS1

The PSSPS1 program processes hospital discharge abstract data, sets comorbidity indicators, and flags inpatient records if they contain outcomes of interest. Outcome indicator names have prefix of "T" (Top). Stratifier variables are constructed at the beginning of the program.

This program assumes that the input data file (consisting of inpatient discharge abstract data) conforms to specific variable names, attributes, and coding conventions. See Table 4 on page 11 for variable names and attributes for the input data file.

Partial definitions of the Patient Safety Indicators are given in Table 1 on page 5. This table is presented to assist those individuals who desire to examine the software source code statements. Complete definitions of the indicators are given in Appendix A of the *"Guide to Patient Safety Indicators"*.

The AHRQ Comorbidity Analysis Program (comoanaly.sas), available at <http://www.ahrq.gov/data/hcup/comorbid.htm>, has been included as part of program PSSPS1. This code adds the comorbidity indicators to each discharge record. The user does not need to access the AHRQ Comorbidity Analysis Program.

Input:

1. User supplied SPSS inpatient data set consisting of administrative hospital discharge abstract data. This data set is a discharge level file with an array of diagnosis and procedure codes, among other data elements.

Output:


1. SPSS dataset containing inpatient records with input variables, stratifiers, comorbidity variables, and flag indicators (TPPSxx and TAPSxx) for the outcomes of interest that will later form the numerators for the PSI rates.
2. Descriptives (with N, MIN, MAX, MEAN and SUM) of all of the numeric variables in the output data file.

Changes:


Change the !LET parameters at the top of the program to specify the following:


1. !INFILE should specify the directory and file name of the input SPSS data set containing discharge records with procedures and diagnoses.
2. !OUTFILE should specify the directory and file name of the output SPSS data set containing patient safety and comorbidity indicators.

3. !PRDAY should specify whether the input SPSS data set contains variables indicating the number of days from admission to the procedure. If the file contains PRDAY variables then !PRDAY should be set to "Yes", otherwise !PRDAY should be set to "No".
4. !MSALEVL is a flag to indicate whether MSAs should be constructed from the county identifiers on the input dataset. Users who wish to calculate rates at the MSA level rather than the county level should set this flag to "Yes", otherwise they should set this flag to "No". The MSA identifier is saved as a separate variable from the county identifier variable, so users who create MSAs will still have the opportunity to construct county and/or MSA level area rates.
5. !LASTDX should specify the number of diagnosis variables on the file. By default this value is set to 30.

 Users with fewer or more diagnosis codes should modify this value accordingly.

6. !LASTPR should specify the number of procedure variables on the file. By default this value is set to 30.

 Users with fewer or more procedure codes should modify this value accordingly.

 *If your variables are named differently from those used in the PSI software (see table 4 on page 11), you MUST either rename those variables prior to running this program or modify PSSPS1 to include the renaming of your variables to match those expected by the software. Similarly, if your variables are not the same type, i.e., character or numeric, as those expected by the software (see Table 4), you MUST modify your variables prior to running this program.*

Program 2: run PSSPSP2

The PSSPSP2 program calculates the observed or raw rates for the hospital-level Patient Safety Indicators, using the data derived in a previous step (PSSPS1). These observed rates are stratified by combinations of hospitals, sex, age, race, and payer categories. The program first totals the indicator flags created by the PSSPS1 program and then for each of the desired stratifiers divides these totals by the hospital discharges in the universe for the indicator. The population denominators are stored in variables with names that have a prefix of "P" (Pop). The Observed rates are stored in variables that have a prefix of "O".

Input:

1. The SPSS dataset that was created in the PSSPS1. This is a discharge-level file that contains inpatient records with input variables, stratifiers, and the 20 new flag indicators for the PSI Hospital outcomes of interest.

Output:

1. SPSS dataset with summary records that contain observed rates (OPPSxx variables where xx refers to the indicator number), the counts of outcomes that formed the numerators of the rates (TPPSxx variables), and the hospital discharge totals that formed the denominators of the observed rates (PPPSxx variables). The output file has records summarized to the various combinations of stratifiers specified in the !STRLEVS parameter that is described in item 5 in "Changes" below.
2. Descriptive statistics (with N, NMISS, MIN, MAX, MEAN, and SUM) of the hospital-level summary records that shows statistics for the OPPSxx observed rates, the TPPSxx counts of outcomes that formed the numerators of the rates, and the PPPSxx counts of outcomes that

formed the denominators of the rates. These means will only be generated if the user included a value of 16 for the !STRLEVS parameter discussed in item 5 in the “Changes” section below.

3. An optional listing of the output summary dataset is provided at the end of the hardcopy printout. This printout may be quite large depending on the number and the types of stratifications that the user requests with the !STRLEVS parameter discussed in item 5 in the “Changes” section below. If the user does not wish to generate this printout, then the !PRINT parameter should be set to “No”.


Changes:

Change the !LET parameters at the top of the program to specify the following:

1. !TEMPDIR should specify the name of an existing directory where SPSS can store temporary files. These files should be deleted once the program completes.
2. !PERMDIR should specify the name of an existing directory where SPSS can store permanent files, including the name of the data set created by program 1.
3. !INDATA should specify the name of the SPSS dataset created by program 1 containing the comorbidity and patient safety indicator flags.
4. !OUTDAT should specify the name of the final SPSS data set that will be created by this program that will contain the provider observed rates.
5. !STRLEVS should specify the levels of stratification for which the program should calculate observed rates. The levels should be specified as a list of numbers (1 to 31 separated by commas) corresponding to the following stratifications shown in Table 5.

Table 5. PSSPSP2 aggregation possibilities					
STRAT	Stratification				
1					Race
2				Payer	
3				Payer * Race	
4			Sex		
5			Sex * Race		
6			Sex * Payer		
7			Sex * Payer * Race		
8		Age			
9		Age			Race
10		Age * Payer			
11		Age * Payer * Race			
12		Age * Sex			
13		Age * Sex * Race			
14		Age * Sex * Payer			
15		Age * Sex * Payer * Race			
16	Hospital				
17	Hospital * Race				
18	Hospital * Payer				
19	Hospital * Payer * Race				

Table 5. PSSPSP2 aggregation possibilities									
STRAT	Stratification								
20	Hospital	*		Sex					
21	Hospital	*		Sex	*			Race	
22	Hospital	*		Sex	*	Payer			
23	Hospital	*		Sex	*	Payer	*	Race	
24	Hospital	*	Age						
25	Hospital	*	Age	*				Race	
26	Hospital	*	Age	*		Payer			
27	Hospital	*	Age	*		Payer	*	Race	
28	Hospital	*	Age	*	Sex				
29	Hospital	*	Age	*	Sex	*		Race	
30	Hospital	*	Age	*	Sex	*	Payer		
31	Hospital	*	Age	*	Sex	*	Payer	*	Race


 STRAT=0 (Overall) is calculated automatically. STRAT=16 (hospital) is the default. For example, if you are using a state inpatient hospital database, you might specify the following STRAT values to the !STRLEVS parameter:


- 0 – provides overall rates for your entire state
- 8 – provides overall rates for your state, broken down by age groups
- 16 – provides rates for hospitals within your state
- 24 – provides rates for age groups within these hospitals

- The !PRINT parameter should indicate whether the final dataset should be listed or not. If !PRINT is set to 'Yes', then the program will generate a listing for each of the 20 hospital indicators, including the numerator, denominator and observed rate for each level of stratification specified by the !STRLEVS parameter. If the user does not want to generate this listing, the !PRINT parameter should be set to 'No'.

Program 3: run PSSPSP3

The PSSPSP3 program calculates age, sex, DRG and comorbidity risk-adjusted rates for each PSI and then calculates smoothed rates.

 *If you are interested in calculating confidence intervals for each PSI, you may request a supplementary module (PSSPSP3CI.SPS) to calculate a 95% CI for each PSI and hospital (see the Interpreting the Results, Confidence Intervals section for additional information). You would run this program in place of PSSPSP3. The user needs to run only PSSPSP3 or PSSPSP3CI, it is not necessary to run both these programs.*

 *When there are fewer than three discharges within a stratification cell, the risk-adjusted rates are set to missing. The smoothed rates that are based on these risk-adjusted rates are also set to missing.*

The output file has records summarized to the various combinations of stratifiers that were specified in the !STRLEVS parameter of the preceding QISPSP2 program. See table 5 on page 18.

Input:

- The discharge-level SPSS data set that was created with the PSSPSP1 program.
- The SPSS data set with summary records that was created with the PSSPSP2 program.
- The AHRQ-supplied text file (COVPSP00.TXT) containing regression coefficients from a regression that was run on a reference SID dataset. These coefficients will be used in the risk

adjustment process. This file is provided to you as part of the PSI module. The text file does not need to be converted to a SPSS file for use with the software.

4. The AHRQ-supplied text file (MSXPSP_0.TXT) containing three arrays for use in the smoothing process. The arrays contain noise estimates, signal estimates, and mean rates for each PSI. This file is provided to you as part of the PSI module. The text file does not need to be converted to a SPSS file for use with the software.

Output:

1. A SPSS dataset with an overall summary record and with hospital-level summary records that contain the three types of indicator rates, along with the components of the initial raw rates. Specifically, the file contains the observed rates (OPPSxx variables), the risk-adjusted rates (RPPSxx variables), the smoothed rates (SPPSxx variables), the counts of outcomes that formed the numerators of the observed rates (TPPSxx variables), and the hospital discharge totals that formed the denominators of the observed rates (PPPSxx variables).
2. Descriptive statistics (with N, NMISS, MIN, MAX, MEAN, and SUM) of the hospital-level summary records that shows statistics for the observed, risk-adjusted and smoothed rates, and statistics for the counts of outcomes that formed the numerators and denominators of the observed rates. These means will only be generated if the user included a value of 16 for the !STRLEV parameter of the preceding QISPSP2 program. See table 5 on page 18.
3. An optional listing of the output summary dataset is provided at the end of the hardcopy printout. This printout may be quite large depending on the number and the types of stratifications that the user requested with the !STRLEVS parameter of the preceding PSSPSP2 program. See table 5 on page 18. If the user does not wish to generate this printout, then the !PRINT parameter should be set to "No".

Changes:

Change the !LET parameters at the top of the program to specify the following:

1. !PERMDIR should specify the name of an existing directory where SPSS can store permanent files. This directory should include the final output data sets created by programs PSSPS1 and PSSPS2P.
2. !TEMPDIR should specify the name of an existing directory where SPSS can store temporary files. These files should be deleted once the program completes.
3. !TEXTDIR should specify the name of the existing directory where AHRQ-supplied text files were saved. The files in this directory should not be converted to SPSS format.
4. !COVARA should specify the name of the AHRQ-supplied text file (contained within the !TEXTDIR directory) where regression coefficients needed to risk adjust the indicators are located.
5. !MSXFILE should specify the name of the AHRQ-supplied text file (contained within the !TEXTDIR directory) containing arrays of coefficients used to smooth the risk adjusted rates.
6. !STRLEVS should specify the stratification levels. This should be consistent with program PSSPSP2.
7. !PRDAY is a flag to indicate whether the user-supplied inpatient discharge records contains fields for number of days from admission to the procedure. If the input file contains PRDAY variables, this flag should be set to "Yes", otherwise it should be set to "No".


7. The !PRINT parameter should indicate whether the final dataset should be listed or not. If !PRINT is set to 'Yes', then the program will generate a listing for each of the 20 hospital indicators, including the numerator, denominator and observed rate for each level of stratification specified by the !STRLEVS parameter. If the user does not want to generate this listing, the !PRINT parameter should be set to 'No'.

Program 4: run PSSPSA2

The PSSPSA2 program calculates the observed or raw rates for the area-level Patient Safety Indicators, using the data derived from Program 1 (PSSPS1). These observed rates are stratified by combinations of area, sex, age, and race categories. The program first totals the six area indicator flags created by the PSSPS1 program, and then for each of the desired stratifiers divides these totals by the pertinent residential population. The population denominators are stored in variables with names that have a prefix of "P" (Pop). The Observed rates are stored in variables that have a prefix of "O".

Input:

1. The SPSS dataset that was created in the PSSPS1. This is a discharge-level file that contains inpatient records with input variables, stratifiers, and the six new flag indicators for the PSI Area outcomes of interest.
2. A text file with Census area residential populations, stratified by area, age, sex, and ethnicity categories. Five such files are currently provided along with the PSI module software. The files are QICTY97, QICTY98, QICTY99, QICTY00, and QICTY01. The user should select the file for the year that best matches the user's discharge datafile.

 *Users do **not** need to convert the ASCII text file to a SPSS dataset for use with the software.*

Output:

1. SPSS dataset with summary records that contain observed rates (OAPSxx variables where xx refers to the indicator number), the counts of outcomes that formed the numerators of the rates (TAPSxx variables), and the residential population totals that formed the denominators of the observed rates (PAPSxx variables). The output file has records summarized to the various combinations of stratifiers specified in the !STRLEVS parameter that is described in item 7 in "Changes" below.
2. Descriptive statistics (with N, NMISS, MIN, MAX, MEAN, and SUM) of the area-level summary records that shows statistics for the OAPSxx observed rates, the TAPSxx counts of outcomes that formed the numerators of the rates, and the PAPSxx residential populations totals. These means will only be generated if the user included a value of 8 for the !STRLEVS parameter discussed in item 5 in the "Changes" section below.
3. An optional listing of the output summary dataset is provided at the end of the hardcopy printout. This printout may be quite large depending on the number and the types of stratifications that the user requests with the !STRLEVS parameter discussed in item 7 in the "Changes" section below. If !PRINT is set to 'Yes', then the program will generate a listing for each of the six area indicators, including the numerator, denominator and observed rate for each level of stratification specified by the !STRLEVS parameter. If the user does not want to generate this listing, the !PRINT parameter should be set to 'No'.

Changes:

Change the !LET parameters at the top of the program to specify the following:

1. !PERMDIR should specify the name of an existing directory where SPSS can store permanent files. This directory should include the final output data set created by programs PSSPS1.
2. !TEMPDIR should specify the name of an existing directory where SPSS can store temporary files. These files should be deleted once the program completes.
3. !TEXTDIR should specify the name of the existing directory where AHRQ-supplied text files were saved. The files in this directory should not be converted to SPSS format.
4. !INDATA should specify the name of the SPSS data set output from program 1 that will be used as input for calculating the observed rates. This data set must be saved in the directory specified by !PERMDIR.
5. !OUTDAT should specify the name of the SPSS data set to be output from this program with the observed rates. This data set will be saved in the directory specified by !PERMDIR.
6. !POPFIL should specify the name of the AHRQ-supplied text file (located within the !TEXTDIR directory) to be used for constructing population denominators used in the area rates.
7. !STRLEVS should specify the comma separated list of stratification levels for which area rates are to be calculated. The stratification levels include the following:


Table 6. PSSPSA2 aggregation possibilities							
STRAT	Stratification						
1						Race	
2				Sex			
3				Sex	*	Race	
4			Age				
5			Age	*		Race	
6			Age	*	Sex		
7			Age	*	Sex	*	Race
8	Area						
9	Area	*				Race	
10	Area	*		Sex			
11	Area	*		Sex	*	Race	
12	Area	*	Age				
13	Area	*	Age	*		Race	
14	Area	*	Age	*	Sex		
15	Area	*	Age	*	Sex	*	Race

🔊 STRAT=0 (Overall) is calculated automatically; STRAT=8 (area) is the default.

For example, if you are using a state inpatient hospital database, you might specify the following !STRLEVS values:

- 0 – provides overall rates for your entire state
- 4 – provides overall rates for your state, broken down by age groups
- 8 – provides rates for areas within your state
- 12 – provides rates for age groups within these areas


8. IMSALEVL is a flag to indicate whether rates in urban areas should be calculated at the MSA level rather than at the county level. Rural areas are calculated at the county level in either case. To calculate urban rates at the MSA level, set this flag to "Yes", otherwise set it to "No".

 Additionally, in order to calculate MSA level area rates, the !MSALEVL flag on program 1 must also have been set to "Yes".

9. The !PRINT parameter should indicate whether the final dataset should be listed or not. If !PRINT is set to 'Yes', then the program will generate a listing for each of the six indicators, including the numerator, denominator and observed rate for each level of stratification specified by the !STRLEVS parameter. If the user does not want to generate this listing, the !PRINT parameter should be set to 'No'.

Reviewing the printed output.

This section contains tips for reviewing some of the printed output from the PSI module. These tips are oriented toward explaining the interrelationships between printout items from different programs and hopefully will help to reveal the nature and structure of the module outputs. For guidance in interpreting the results, see the next section which begins on page 27.

 *The final output of the PSI module is a LISTING from PSSPSP3 (for hospital-level indicators) and PSSPSA2 (for area-level indicators). All interim printouts are for checking and trouble-shooting. However, if you have elected to refrain from risk adjusting and smoothing the hospital indicators, then the final hospital indicator output will be the LISTING at the end of the PSSPSP2 program. Risk-adjusted and smoothed rates are not calculated for the area-level PSIs.*

PSSPS1

The initial printout from the PSSPS1 program contains descriptive statistics output for all of the numeric variables (including the comorbidity indicators) in the output discharge-level dataset. It will contain information for the newly constructed TAPSxx and TPPSxx flag variables that will later form the numerators for the indicator rates. For each TAPSxx and TPPSxx flag variable:


- The SUM will contain the total number of observations in the dataset that were found to have the particular outcome of interest.
- The Hospital indicator numerators (TPPS01-TPPS20) and the Area indicator numerators (TAPS21-TAPS26) will all have a value of one, zero, or missing ('.'). For these indicators, a value of zero was assigned to the TAPSxx/TPPSxx flag if a particular observation was part of the population at risk for a particular indicator but did not have the particular outcome of interest. For example, TPPS20 = 1 indicates a patient who had an obstetric trauma from a Cesarean section, while TPPS20 = 0 identifies a patient who had a C-section but no trauma complication.
- The MEANs will contain a close approximation of the eventual overall observed indicator rates. The values will change slightly after PSSPSA2/PSSPSP2 has applied additional parts of the indicator definitions.
- N lists the number of observations in the dataset with non-missing values. N will contain the denominator for the observed indicator rate.

PSSPSP2

The purpose of this printout is to provide the observed rates for all hospital-level indicators. The default printout shows results by hospital, but other levels of output (as shown in Table 5) can be specified as well. Note that risk-adjusted and smoothed rates are shown in the output from the next program.

The printout from the PSSPSP2 program contains descriptive statistics on the hospital-level summary records for all of the numeric variables in the output summary dataset. It will contain information for the newly constructed OPPSxx rates, the PPPSxx denominators, and the TPPSxx numerators.


- The STRAT variable described in the descriptives table identifies the stratification level for the records in the output data set. The STRAT variable corresponds to the !STRLEVS parameter values that were specified (see item 5 in “Changes” under the PSSPSP2 program). In this case, STRAT always assumes the value of 16, since only the hospital-level records are selected.
- The N statistic for STRAT contains the number of records in the output summary data set. A TPPSxx numerator variable with a value for N that is lower than the N value for STRAT indicates that there were no outcomes of interest in some of stratification cells. Similarly, a PPPSxx denominator variable with a lower value for N than STRAT indicates that for some stratification cells, there were no hospital discharges with the outcome of interest.
- The MINIMUM value for the TPPS01-TPPS20 numerators will be zero since values of zero were assigned for observations that were part of the population for the rate denominator but did not have the particular outcome of interest to be included in the rate numerator.

 For the observed rates, the User should **NOT** report the MEANS displayed here, but instead should refer to the overall means in the output dataset that are also present in the subsequent LISTING. The MEANS given here are means of hospital means; i.e., all hospitals are given equal weight, regardless of the number of discharges in the different hospitals.

- The SUMs for the counter variables TPPSxx and PPPSxx will contain overall file totals for these data elements. The SUMs for the observed rates have no intuitive meaning.

If the !PRINT parameter is set to “Yes”, the subsequent LISTING output contains a complete dump of the output summary file. Listed for each record are the stratification values, the PPPSxx population denominators for the 20 indicators, the TPPSxx outcome numerators for the 20 hospital indicators, and the 20 OPPSxx observed rates.

PSSPSP3

 The LISTING at the end of this program provides your **final output**. (This printout appears if the !PRINT parameter is set to “Yes”—the default setting.) It lists the numerator and the denominator, along with the observed, risk-adjusted, and smoothed rates for all indicators, as shown in the following example.

STRAT	HOSPID	AGECAT	SEXCAT	PAYCAT	RACECAT	TPPS04	PPPS04	OPPS04	RPPS04	SPPS04
0	1103	7322	.150642	.184167	.183908
16	67000	0	8	.000000	.097114	.166435
16	71000	13	85	.152941	.225124	.191035
16	72000	0	6	.000000	.117922	.168236
16	76000	27	225	.120000	.162791	.165709
16	81000	52	220	.236364	.228673	.205974
16	85000	21	159	.132075	.181800	.176450
16	87000	1	19	.052632	.087305	.160480
16	89000	3	23	.130435	.226091	.178231
16	91000	21	150	.140000	.187632	.179297
16	93000	40	279	.143369	.180632	.177203
16	95000	39	219	.178082	.226401	.204526
16	96000	14	82	.170732	.194382	.179237
16	104000	0	.	.	.
16	106000	4	23	.173913	.137757	.165792
16	108000	7	65	.107692	.175852	.172117
16	109000	33	223	.147982	.188629	.181584
16	111000	13	79	.164557	.166927	.169141
16	114000	71	420	.169048	.177468	.175694
16	115000	2	14	.142857	.158940	.169349
16	117000	36	452	.079646	.144308	.150487
16	118000	173	1276	.135580	.170824	.170781
16	119000	42	307	.136808	.175408	.173833
16	121000	10	107	.093458	.134601	.154905
16	124000	161	906	.177704	.205327	.200641
16	125000	9	80	.112500	.179932	.173853
16	128000	1	3	.333333	.351560	.174179
16	129000	49	282	.173759	.199062	.189534
16	130000	67	397	.168766	.202402	.194041
16	134000	52	256	.203125	.213217	.198052
16	137000	3	73	.041096	.098798	.145890
16	140000	22	164	.134146	.167299	.168723
16	141000	23	98	.234694	.256188	.205668
16	142000	60	336	.178571	.190744	.184747
16	143000	23	225	.102222	.149181	.157327
16	144000	11	61	.180328	.214140	.183644

Number of cases read: 36 Number of cases listed: 36


In this example printout, record 1 (with HOSPID = .) is the overall average for the entire dataset (STRAT = 0). The remaining observations are individual hospitals, or hospitals (STRAT = 16). Data will appear under the column headers, AGECAT, SEXCAT, PAYCAT, and RACECAT, if you specify these levels of aggregation (shown in Table 5).

You may wish to express the results in more understandable terms:

- Multiply the rates by 100 to express them as a percentage (e.g., .150642 * 100 = 15.06%)
- To create standardized rates or rates per 1,000 discharges, multiple the output by 1,000.

The descriptive statistics table is generated from the permanent hospital-based output file and is provided for trouble-shooting and understanding your data, but not for reporting. The means are provided just for the hospital level summary records.

- The N for all but the TPPSxx counter variables should contain the number of hospitals in the users database.
- The MEANS, MINIMUMs, and MAXIMUMs have their normal meaning and provide comparisons among the different hospitals or hospitals in your database.


 *For the (observed, risk-adjusted, or smoothed) rates, the User should **NOT** report the MEANS displayed here, but instead refer to the overall means in the output dataset in the PROC PRINT labeled "FINAL OUTPUT." The MEANS given here are means of hospital means; i.e., all hospitals are given equal weight, regardless of the number of discharges in the different hospitals.*


- If the !PRINT parameter is set to "Yes", at the end of the printout is a LISTING on the final hospital-based output file. Listed for each level of stratification are the stratifiers, the numerator and denominator totals for the observed rates, and the observed, risk-adjusted, and smoothed rates.

PSSPSA2

The purpose of this printout is to provide the observed rates for all area-level indicators. The default printout shows overall results and results by area, but other levels of output (as shown in Table 6) can be specified as well.

- The printout from the PSSPSA2 program contains descriptives for all of the numeric variables in the output summary dataset. It will contain information for the newly constructed OAPSxx rates, the PAPSxx denominators, and the TAPSxx numerators. The MEANS table is generated from the permanent area-based output file. The means are provided just for the area-level summary records.
- The STRAT variable described in the first row of the descriptives table identifies the stratification level for the records in the output data set. The STRAT variable corresponds to the !STRLEVS parameter values that were specified (see item 6 in “Changes” under the PSSPSA2 program). In this case, STRAT always assumes the value of 8 since only the area-level records are selected.
- The N statistic for STRAT contains the number of records in the output summary data set. A TAPSxx numerator variable with a lower value for N than STRAT indicates that there were no outcomes of interest in some of stratification cells. Similarly, a PAPSxx denominator variable with a lower value for N than STRAT indicates that for some stratification cells, the Census residential population estimate was zero.
- The N for all but the TAPSxx counter variables should contain the number of different areas (MSAs and counties) in your database.
- The MEANS, MINIMUMs, and MAXIMUMs have their normal meaning and provide comparisons among the different areas in your database.
- The MINIMUM value for the TAPSxx numerators will be one or higher.
- The SUMs for the counter variables TAPSxx and PAPSxx will contain overall file totals for these data elements. The SUMs for the observed rates have no intuitive meaning.

 *For the observed rates, the User should **NOT** report the MEANS displayed here, but instead should refer to the overall means in the output dataset that are also present in the subsequent LISTING. The MEANS given here are means of the area means; i.e., all areas are given equal weight, regardless of the population in the different areas.*

 *The LISTING at the end of this program provides your **final output**. (This printout appears if the !PRINT parameter is set to “Yes”—the default setting.) It lists the numerator and the denominator, along with the observed, risk-adjusted, and smoothed rates for all indicators, as shown in the example output below.*

MSA	AGECAT	SEXCAT	RACECAT	TAPS25	PAPS25	OAPS25
.				661	20953241	.000032
40				18	126441	.000142
320				0	218321	.000000
640				2	1263559	.000002
840				12	384977	.000031
1240				4	336991	.000012
1260				10	152657	.000066
1880				1	380689	.000003
1922				85	5254221	.000016
2320				45	682111	.000066
3362				219	4693524	.000047
3810				0	314252	.000000
4080				0	194868	.000000
4420				1	208797	.000005
4600				42	242882	.000173
4880				9	573920	.000016
5800				0	236345	.000000
7200				0	103970	.000000
7240				55	1599378	.000034
7640				0	110988	.000000
8360				4	129762	.000031
8640				0	175418	.000000
8750				1	84075	.000012
8800				1	214044	.000005
9080				0	140326	.000000
48001				0	55119	.000000
48003				0	12944	.000000
48005				14	80270	.000174
48007				0	22548	.000000

Record 1 (with MSA = '.') is the overall average for the entire dataset (STRAT = 0). The remaining observations are individual areas (STRAT = 8).

You may wish to express the results in more understandable terms:

- Multiply the rates by 100,000 to express them as a rate per 100,000 population (e.g., .000032 * 100,000 = 3.2 accidental puncture/lacerations per 100,000 population).

Interpreting the results

After performing all of the steps as previously outlined, users will have a set of observed estimates of performance for each Area PSI and three estimates of performance for each Hospital PSI – observed, risk-adjusted, and MSX smoothed (i.e., reliability-adjusted) estimates. This section will outline how the different types of estimates can be used.

The Patient Safety Indicators are not intended as definitive quality measures. Many factors can and do influence performance on quality indicators. Some of these factors are independent of quality. From our empirical analyses we estimated that generally less than 1 percent (range 0.2-15%) of total patient level variation for each QI is attributable to actual hospital-level differences. Though this number appears low, we also estimated that often most of this remaining hospital-level variation reflects true signal; that is, the observed performance is not likely the result of random variation. We also noted that these true differences in hospital performance are potentially clinically important, in that they involve severe and potentially preventable complications and adverse events. Thus, though these indicators are not intended and should not be used as definitive measures of quality, they are useful low-cost screens that can potentially illuminate differences between hospitals that may be addressable using quality and patient safety improvement techniques.

While very high rates may indicate quality problems, it is often best to compare hospital rates with other similar hospitals. These “peer groups” would ideally be as similar as possible in potentially important factors, such as case mix, socioeconomic status and other demographics. Similar hospitals should perform similarly on PSIs. Potential characteristics that have been used to define peer groups have been teaching status, bed capacity, socioeconomic status of the population, and urban or rural location.

Performance on a single PSI often cannot reliably indicate actual quality differences. Examining all the indicators together is likely to produce a more complete picture of overall quality of care. Hospitals that perform better (or worse) than average on one measure tend to perform better (or worse) on others. In other words, performance on indicators is often related, and performance on multiple similar indicators could also be examined as summary indicators for an overall picture of quality. Similarly, indicators may be related across indicator types. *Groups of indicators have not been validated as sets, but using them as such may be useful for exploration purposes.*

Observed means.

The interpretation of the observed (raw) PSI rates may be complicated by concerns over noise and bias. However, the observed rates do not require complex statistical methods as do the risk-adjusted and MSX smoothed rates. For this reason, they are the most intuitive estimates of PSI performance. Observed rates can be used as a baseline measure when comparing to risk-adjusted and MSX smoothed rates, to determine the impact of risk adjustment/smoothing.


Observed rates may fluctuate for some hospitals from year to year. While some of these fluctuations may indicate true differences in performance, others are due only to random chance (i.e., non-systematic differences among hospitals in patient characteristics that are not observed, and therefore cannot be adjusted for). For this reason, it is often beneficial to examine observed rates for a given hospital over several consecutive time periods. This will help to diminish the chance of inaccurately judging poorer than average or better than average performance during one year as true performance.

Risk-adjusted rates. Hospital PSIs only.

Simply put, risk-adjusted rates are the estimated performance of hospitals on the PSIs if those hospitals had an 'average' case mix. This average case mix is estimated and included in the software using estimates from 29 States in the SID databases. Estimates of the average case mix do not reflect only the hospitals in the dataset being analyzed, but rather the distribution in age, sex, DRG and co-morbidity categories among the hospitals in the SID data.

Most hospital indicators are risk adjusted using age, sex, DRGs and co-morbidities. A few hospital indicators cannot be adjusted using DRGs since the population at risk is characterized by a few DRGs. These are the OB related indicators.

For Transfusion reaction (PSI #16) the outcome of interest is a relatively rare event and the observed rates will be extremely low. No risk-adjustment or smoothing is performed for this indicator.

 *The output will not include any risk-adjusted or smoothed rates for PSI #16, Transfusion reaction. The software will generate observed rates only.*

Users should compare risk-adjusted rates with the observed rates. This will indicate the impact of risk adjustment and presumably case mix on indicator performance. Hospitals with large changes in performance between observed and risk-adjusted rates presumably have a more or less complex case mix.

Additional information on the risk-adjustment methodology is available in the *Guide to Patient Safety Indicators*. In summary, all indicators were examined empirically by developing and conducting statistical tests for precision, bias, and relatedness of indicators. For each indicator, the project team calculated five different estimates of hospital performance:

1. The raw indicator rate was calculated using the number of adverse events in the numerator divided by the number of discharges in the population at risk by hospital. For the area indicators, the denominator is the population of the Metropolitan Statistical Area (MSA), New England County Metropolitan Area (for the New England States) or county (for non-MSA areas) of the hospital.

2. The raw indicator was adjusted using a logistic regression to account for differences among hospitals (and areas) in demographics (specifically, age and gender). Age was modeled using a set of dummy variables to represent 10-year categories except for young children, whose age categories are narrower (i.e., less than 1, 1-4, 5-14, 15-24, 25-34, 35-44, 45-54, 55-64, 65-74, 75-84, and 85 or more years), along with a parallel set of age-gender interactions. Because of sparse cells, certain age categories were combined or omitted for selected indicators, such as the obstetric indicators.
3. The raw indicator was adjusted to account for differences among hospitals in age, gender and modified DRG category.
4. The raw indicator was adjusted to account for differences among hospitals in age, gender, modified DRG, and comorbidities (defined using an adaptation of the AHRQ comorbidity software) of patients.
5. Multivariate signal extraction (MSX) methods were applied to adjust for reliability by estimating the amount of “noise” (i.e., variation due to random error) relative to the amount of “signal” (i.e., systematic variation in hospital performance or the ‘reliability’) for each indicator. This or similar “reliability adjustment” has been used in the literature for similar purposes. Multivariate methods (taking into account correlations among indicators to extract additional signal) were applied to most of the accepted indicators. The exceptions were Death in Low Mortality DRGs and Failure to Rescue. Only univariate signal extraction methods (smoothing) were applied to these two indicators and to the experimental indicators, because these indicators possibly cover broader clinical concepts. Correlations between these indicators and other indicators may not reflect correlations due to quality of care, and thus inclusion of these indicators may adversely affect the MSX approximations.

In our risk-adjustment models, we calculated hospital fixed effects using the standard method with logistic models of first estimating the predicted value for each discharge, then subtracting the actual outcome from the predicted, and averaging the difference for each hospital (and adding in the overall rate) to get the hospital fixed effect estimate. The logistic model was appropriate given the rare occurrence of many of the PSIs and the dichotomous outcomes. The risk-adjustment model for each PSI accounts for patient demographics (age, sex) and DRG/co-morbidity category. The risk-adjusted rates reflect the estimated performance on the PSI for each hospital if that hospital had the ‘average’ case mix among all hospitals in the estimation sample.

Second, because the fixed effects are calculated indirectly with a logistic model, the estimated risk-adjusted rates may be negative (that is, the model does not require the rates to be between zero and one). This might happen, for example, if the observed rate is close to zero, and the hospital has a more severe than average case mix. The model would predict that the hospital’s performance with an ‘average’ case mix would be less and subtract some adjusted amount from the observed rate. If that observed rate is already nearly zero, the adjusted rate may be negative. A negative rate means that the observed rate is close to zero, and that the hospital has a less-severe-than-average case mix. In order to facilitate the interpretation of the estimates, the PSI module imposes a lower bound of zero; i.e., negative values are set to zero.

MSX smoothed rates. Hospital PSIs only.

The PSSPSP3 program also yields the estimated MSX smoothed (i.e., reliability-adjusted) rates for each Hospital PSI. These smoothed rates are estimated using multivariate signal extraction (MSX). MSX is a Bayesian smoothing technique. Generally, indicators defined on relatively small populations per hospital or for relatively rare events are very noisy measures. In other words, many other factors other than quality can influence the observed rate. For factors that we can observe, such as the patient’s age or gender or DRG, risk adjustment can account for their influence. There are many other clinical and non-clinical factors that we cannot observe. However, the MSX method estimates how much of an impact random differences in these factors have on the observed rate. The SPSS program accomplishes this by adjusting toward the overall mean according to the degree of reliability of the specific hospital indicator.

The MSX method estimates reliability by looking at hospital performance among multiple indicators and across multiple years for a single indicator. If hospitals that perform high in one year also tend to perform high in the next, and hospitals that are low in one indicator tend to perform low in other related indicators, then the methods incorporate this information in forming an estimate of how much of a hospital's performance on a single indicator in a single year is "signal" and how much is "noise". PSSPSP3 reports the "signal" estimate, after removing the "noise".

These estimates are referred to as "smoothed" estimates because when the raw data is reported annually on a graph, for example, the line connecting the points may show massive fluctuations, seemingly with no apparent pattern. Some of these fluctuations are due to random differences in factors that influence performance, so once the impact of these factors is removed, fluctuations are reduced, or smoothed out. Improvements in performance will still be observed, however, they will only be observed when it is likely that these improvements are actually true differences, and not simply random variation.

The estimates of how related the Patient Safety Indicators are among measures and over time for a typical hospital are based on estimates using the 29 states in the SID database. Therefore, PSSPSP3 does not require multiple hospitals, measures, and years of data to operate. Instead shrinkage estimates were calculated from the combination of 29 statewide HCUP datasets to which the MSX had been applied. These shrinkage estimates are the best estimates of the impact that MSX smoothing would have on a particular PSI if MSX could be applied to the dataset being analyzed. These shrinkage estimates are then applied to the adjusted rates for each PSI, giving 'smoothed' estimates of the rates.

The smoothed estimates represent the best guess of what the hospital performance would be for any given PSI if the data contained only a minimal amount of noise. This has several implications. First, users can be more certain that changes observed over time represent true differences in performance, rather than random variation. Second, in any given year, users can be more certain that the rate does not reflect only random noise. This can best be seen by comparing the smoothed estimates with the observed raw rates, which will indicate the impact of MSX on that indicator. A large impact means that the observed rate of the indicator is less precise, meaning it contains more noise relative to other indicators sustaining less impact. Low precision can result from small numbers of observations, or for relatively rare events. Therefore, the impact of the MSX estimate will in general be greater for small hospitals and for infrequent events. Plotting the MSX estimates over time will reflect more persistent performance from year-to-year, and the correlations among indicators will generally be stronger.

Comparing observed, risk-adjusted and smoothed rates. Hospital PSIs only.

The purpose of the analysis determines which rates the user should look at in evaluating the performance of a hospital. If the user's primary interest is to focus on a particular hospital without any comparisons to other hospitals, simply examine the overall observed rate for the entire hospital, as well as further breakdowns by age, sex, payer, and race/ethnicity which are the output created by the PSSPSP2 programs. Trend data are recommended for identifying changes over years.

If the purpose of the analysis is to compare the performance of a particular hospital with national, state, or regional averages or performance of other selected hospitals, then both observed and risk-adjusted rates should be examined. Variation in observed rates across hospitals is attributable to a variety of factors including differences in patient case mix or population demographics, disparity in access to and quality of care, and other hospital characteristics ('systematic factors'), and random factors (non-systematic factors or 'noise'). Comparing observed and risk-adjusted rates can reveal if there is any difference between the hospital's population and the population of other hospitals. If the difference is minimal, one can compare the observed rate with the overall average across all hospitals. However, to account for differences in patient case mix or population demographics among different hospitals, risk-adjusted rates should be used for hospital comparisons.

After removing the impact of patient case mix or population demographics, if users want to examine potential existence of random factors, comparisons can be made between the risk-adjusted rates and the smoothed rates to determine if the differences in risk-adjusted rates across hospitals are due to

systematic, as opposed to random, factors. The following two tables provide guidance on how to interpret differences in observed, risk-adjusted, and smoothed rates.

Comparing Observed Rates with Risk-adjusted Rates

Purpose: To identify if there is any difference in the patient case mix of the hospital relative to that of all hospitals combined

Observed rate > Risk-adjusted rate	The hospital's patient population for the condition or procedure has a higher risk of complication/adverse event due to its case mix (for example, older or a greater proportion of a higher-risk DRG/ comorbidity).
Observed rate < Risk-adjusted rate	The hospital's patient population for the condition or procedure has a lower risk of complication/adverse event due to its case mix (for example, younger or a greater proportion of a lower-risk DRG/comorbidity).
Observed rate = Risk-adjusted rate	Mortality indicators: The hospital's patient case mix for the condition or procedure is similar to other hospitals', suggesting that patient composition is not a contributing factor to the hospital's performance for the complication/adverse event indicator.

Comparing Risk-adjusted Rates with Smoothed Rates (optional)

Purpose: To determine if differences in risk-adjusted rates across hospitals are due to systematic or random factors after patient case mix or demographic composition is taken into account.

Risk-adjusted rate similar to Smoothed rate	The hospital's risk-adjusted rate is more precise, an accurate reflection of the hospital's systematic ('true') rate. If there is any difference between the risk-adjusted rate and the overall average ¹ , such difference is likely to be due to systematic factors.
Risk-adjusted rate different from Smoothed rate	The hospital's risk-adjusted rate is less precise ² ; that is, some of the difference between the risk-adjusted rate and the overall average is due to random ('noise') factors. The smoothed rate is thus a more accurate estimate of the systematic rate than the risk-adjusted rate. Comparing the smoothed rate to the overall hospital average reveals more accurately the 'true' performance of the hospital.

Notes: 1. The overall hospital average refers to the weighted average risk-adjusted rates across all hospitals.

2. The risk-adjusted rates for hospitals with small population are generally less precise.

Confidence Intervals.

The software does not include a procedure for calculating confidence intervals for the PSI. Any such procedure should be evaluated for robustness depending on the hospital's denominator and rate. User's may request a supplementary module (PSSPSP3CI.SPS) to calculate a 95% CI for each PSI and hospital using a procedure based on a normal approximation that is generally appropriate. The module can be requested by sending an e-mail to support@qualityindicators.ahrq.gov and specify "PSI Confidence Interval Syntax Request" in the subject line and indicate in the body of the e-mail you would like the SPSS syntax.

Benchmark timings

The benchmark runtimes given below are from runs made on a Microsoft Windows 98 PC, with an Intel Pentium III, 1.2 GHz processor, and 512MB of RAM. The dataset used contained 300,000 records with 10 diagnosis fields and 6 procedure fields.

Step	Run time (in seconds)
Run PSSPS1	1,980
Run PSSPSP2	29
Run PSSPSP3	1,578
Run PSSPSA2	23

User Support and Information

We would like to receive your feedback on the AHRQ Quality Indicators.

Our Internet address for user feedback is: support@qualityindicators.ahrq.gov

We offer a listserv to keep you informed on the Quality Indicators. We encourage you to sign up for this free service. All you need is a computer, Internet access, and an E-mail address. It works just like other electronic distribution lists.

Here's how to register:

1. Send an E-mail message to: listserv@qualityindicators.ahrq.gov.
2. On the subject line, type: Subscribe. For example:
Subscribe
3. In the body of the message type: sub Quality_Indicators-L and your full name. For example:
sub Quality_Indicators-L John Doe
4. You will receive a message confirming that you are signed up.

If you have any questions, contact AHRQ QI Support at the e-mail noted above. You should receive an automatic response e-mail message confirming receipt of your e-mail. If you do not receive a confirmation message, call (888) 512-6090.